Appl Intell (2009) 31: 31–46
DOI 10.1007/s10489-007-0110-y

Quantum minimization for adapting ANFIS outputs to its nonlinear generalized autoregressive conditional heteroscedasticity

Bao Rong Chang · Hsiu-Fen Tsai

Published online: 22 December 2007
© Springer Science+Business Media, LLC 2007

Abstract Volatility clustering degrades the efficiency and effectiveness of time series prediction and gives rise to large residual errors. This is because volatility clustering suggests a time series where successive disturbances, even if uncorrelated, are yet serially dependent. Traditional time-series forecast model such as grey model (GM) or auto-regressive moving-average (ARMA) has often encountered the overshoot effect, thus leading to the deterioration of its predictive accuracy. To overcome the overshoot and volatility clustering problems at the same time, an adaptive neuro-fuzzy inference system (ANFIS) is combined with a nonlinear generalized autoregressive conditional heteroscedasticity (NGARCH) model that is adapted by quantum minimization (QM) so as to tackle the problem of overshooting situation and time-varying conditional variance residual errors. The proposed method significantly reduces large residual errors in forecasts because the overshoot and volatility clustering effects are regulated to trivial levels. Two experiments using real financial and geographic data series, respectively, compare the proposed method and a number of well-known alternative methods. Results show that forecasting performance by the proposed method produces superior results, with good speed of computation. Goodness of fit of the proposed method is tested by Ljung-Box Q-test. It is concluded that the ANFIS/NGARCH composite model adapted by QM performs very well for improved predictive accuracy of irregular non-periodic short-term time series forecast and will be of interest to the science of statistical prediction of time series.

Keywords Time-series prediction · ANFIS/NGARCH composite model · Quantum minimization · Overshoot effect · Volatility clustering · Nonlinear heteroscedasticity

B.R. Chang (✉)
Department of Computer Science and Information Engineering, National Taitung University, Taitung, Taiwan 950
e-mail: brchang@nttu.edu.tw

H.-F. Tsai
Department of International Business, Sue-Te University, Kaohsiung, Taiwan 824

1 Introduction

Time-series forecasting uses a set of historical values, i.e. a "time series," to predict future values. Time series values can represent any data series that is distributed over time, for example monthly air temperature, daily electricity consumption or hourly stock volume. A prediction model based on a long history of observations is considered the best way to forecast long-run trends in a long-term and reasonably consistent time series. It is much more difficult to construct predictors that are capable of forecasting non-periodic short-term time series [1]. In practice, predictions are obtained by forecasting a value at the next time instant based on a prediction algorithm [1]. The autoregressive moving-average (ARMA) is a traditional method very suitable for forecasting regular periodic data like seasonal or cyclical time series [2]. On the other hand, ARMA does not work well on irregular or non-periodic data sequences such as international stock prices or future volume indices [3]. This is because ARMA lacks a learning mechanism and cannot tackle large fluctuations in a complex time series.

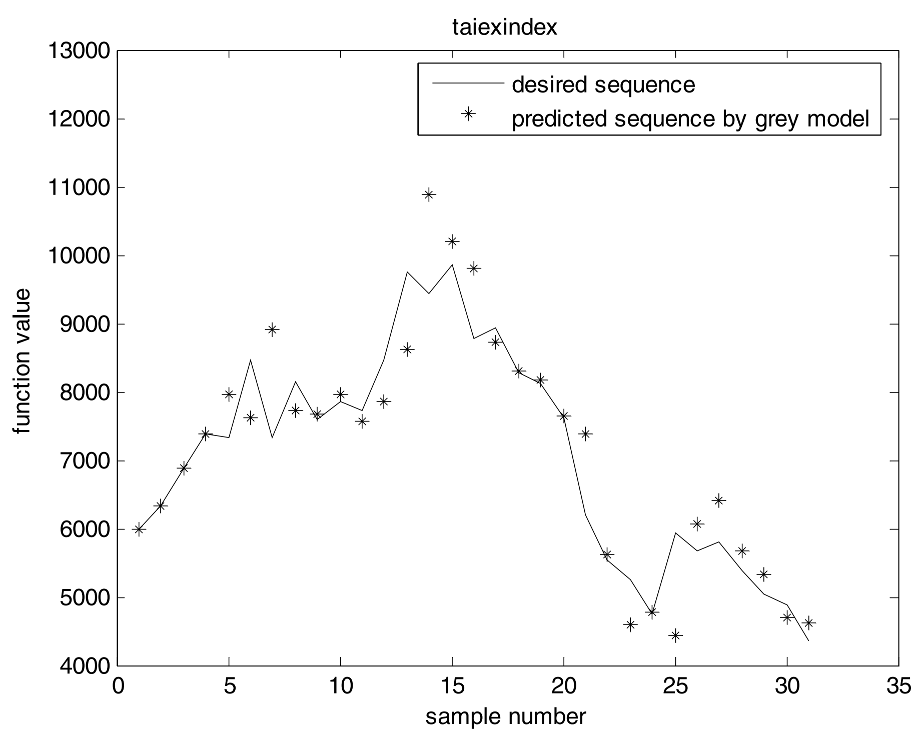
The mathematical models associated with traditional forecasting methods are linear and fail when the data they model is highly nonlinear. In order to better model irregular dynamic behavior, new intelligent methodologies such as artificial neural networks, knowledge-based systems and genetic algorithms have attracted attention. In particular, neural networks are finding extensive use for financial forecasting. For example, the back-propagation neural network (BPNN) [4] and radial basis function neural network (RBFNN) [5] have been successfully applied to time series forecasting but require a large amount of pattern/target training data to capture the dynamics of the time series. Providing a BPNN or RBFNN with insufficient (scarce) training data leads to premature completion of training typically characterized by over-fit or under-fit, producing a system which cannot guarantee adequate prediction results with a specified level of accuracy.

An alternate predictor, the grey model [6], has been widely applied to non-periodic short-term forecasts and its performance on time series prediction is better than Holt-Winters smoothing [2] or ARMA. However, grey model commonly encounters the overshoot phenomenon [1] whereby huge residual errors emerge at the inflection points of a data sequence, i.e. when the slope changes from positive to negative and vice versa. For example, Fig. 1 compares the actual monthly TAIEX stock price index for 31 months (January 1999 to July 2001) with the grey model prediction. The grey model predictions (marked by "∗") reveal the overshoot problem at the turning point regions of sample numbers 7, 14, 25, 26 and 27. Clearly, the overshoot phenomenon seriously weakens grey model prediction accuracy.

The adaptive neuro-fuzzy inference system (ANFIS) [7] has been widely applied to random data sequences with highly irregular dynamics [8], e.g. forecasting non-periodic short-term stock prices [1]. The success of ANFIS can be attributed to two factors: (a) the designated distributive inferences stored in the rule base; (b) the effective learning algorithm for adapting the system's parameters [7]. ANFIS is a Sugeno-type fuzzy inference system in which the parameters associated with specific membership functions are adaptable using either a back propagation gradient descent algorithm alone or in combination with a least squares method [9]. This allows the fuzzy system to quickly fit a time series that might be a non-periodic short-term data sequence. However, volatility clustering effects [10] in the data sequence prevent ANFIS from reaching desired levels of accuracy. This is because ANFIS itself cannot resolve conditional heteroscedasticity [10] (non-constancy of the variance of a measure, i.e. time-varying variance) in residual errors within a data sequence to overcome volatility clustering effects. In application to non-periodic short-term forecast, ANFIS predictions make large residual errors due to high residual variance, consequently degrading prediction accuracy. Obviously, whether or not time-varying conditional variance [11] in residuals is resolved, conditional heteroscedasticity significantly affects a model's goodness of fit. Thus, how to deal with the problem of volatility clustering is an important and interesting issue.

A model called generalized autoregressive conditional heteroscedasticity (GARCH) [11] was introduced to capture the changes in time-varying conditional variance and improve the generalization or stability of the forecast. Further, the same study presented a revised version called a nonlinear generalized autoregressive conditional heteroscedasticity (NGARCH) [11] for resolving volatility clustering ef-

Fig. 1 Overshoot phenomenon in grey model prediction



fects. In our present study, a scheme is proposed to remedy situations of nonlinear conditional heteroscedasticity by incorporating NGARCH into an ANFIS system, which will be denoted below as the ANFIS/NGARCH composite model. To do so, an adaptation called quantum minimization (QM) [12] is applied to tuning the coefficients of a linear combination of ANFIS and NGARCH so that large residual error is compensated by NGARCH and near-optimal solutions can be obtained.

2 ANFIS/NGARCH composite model resolving volatility clustering

2.1 Volatility clustering problem

Volatility commonly means the risk or uncertainty measure associated with a financial time series and is generally associated with the standard deviation of that time series. A certain characteristic commonly associated with financial time series is called volatility clustering, in which large changes tend to follow large changes, and small changes tend to follow small changes. In either case, the changes from one period to the next are typically of an unpredictable sign. Volatility clustering, or persistence, suggests a time series model where successive disturbances, even if uncorrelated, are yet serially dependent. Thus large disturbances, either positive or negative, can be used to construct the variance forecast of the next period's disturbance based on historical information about volatility clustering. For example, the backward-difference values of the monthly equity volume index futures at the London derivatives market for 48 months (Jan. 1999 to Dec. 2002) are shown in Fig. 2 with its zero mean value (solid line). Volatility clustering can be seen in this plot, for example clusters of small changes occur around samples number 1 to 15 and 36 to 38, while clusters of large changes occur around samples number 16 to 36 and 39 to 48. The statistical methodology called Generalized Autoregressive Conditional Heteroscedasticity (GARCH) is well suited to dealing with the problem of volatility clustering in financial time series. Generally speaking, we treat heteroscedasticity as time-varying variance, i.e. volatility. The word conditional suggests a dependence on the observations of the immediate past, while autoregressive represents a feedback operation to incorporate past observations into the present. In other words, GARCH includes past variances in the explanation of future variances and allows users to model the serial dependence of volatility. Conventionally, GARCH is formulated as a linear function of the residuals themselves and the conditional variance of residuals. However, by considering the capabilities of the type of model and matching it to the reality being modeled, then for more accurate estimation or prediction of time-varying conditional variances of residuals, a nonlinear type of GARCH is best taken into account. Since our goal is the handling of volatility clustering in nonlinear time series, the nonlinear GARCH (NGARCH) is considered the superior approach and will be discussed in the following section.

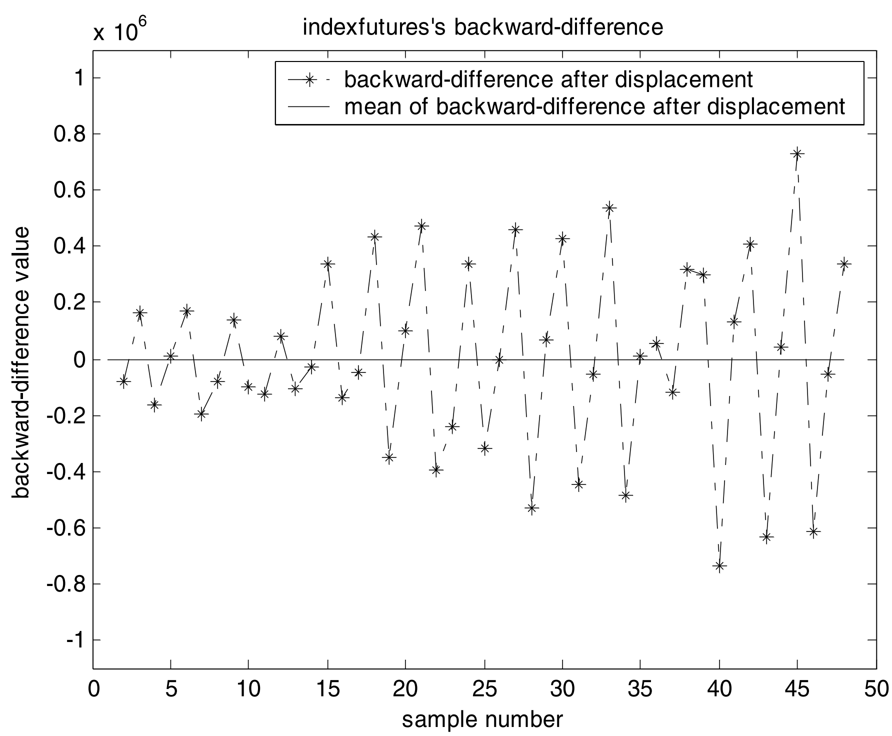


Fig. 2 Data sequence shows volatility clustering effect

2.2 NGARCH resolving volatility clustering

Generally speaking, a time series can be viewed as a sequence of random observations that may reveal some degree of correlation from one observation to the next. We can use this correlation structure to predict the future values of a random sequence on the basis of the historical observations. To analyze the correlation structure, it is necessary to decompose the time series into a deterministic component (i.e. the forecast) and a random component (i.e. the residual error associated with the forecast) [13]. To represent an observed time series, this decomposition can be expressed as in the equation

$$y(t) = f(t-1, \mathbf{x}) + e_{\text{resid}}(t), \tag{1}$$

where a deterministic component $f(t-1, \mathbf{x})$ represents the forecast of the current response as a function of any information known at time $t-1$, including past residuals $\{e_{\text{resid}}(t-1), e_{\text{resid}}(t-2), \ldots\}$, past observations $\{y(t-1), y(t-2), \ldots\}$ and any other relevant explanatory time series data vector $\mathbf{x}$, while a random component $e_{\text{resid}}(t)$ stands for the residual error in the mean of $y(t)$. The random disturbance $e_{\text{resid}}(t)$ represents single-period-ahead forecast error.

In order to explain how random disturbance $e_{\text{resid}}(t)$ is associated with time-varying conditional variances, an ordinary ARMAX/NGARCH composite model [11] is presented to demonstrate its structure. Note that this composite method consists of a conditional mean model and a variance model, with ARMAX representing the conditional mean model and NGARCH representing the conditional variance model. In contrast with (1), a deterministic component $f(t-1, \mathbf{x})$ is calculated from ARMAX and a random component $e_{\text{resid}}(t)$ is determined by NGARCH.

ARMAX(r, m, N_x) encompasses autoregressive (AR), moving-average (MA) and regression (X) models, in any combination, as expressed below

$$\begin{aligned} y_{\text{armax}}(t) &= C^{\text{armax}} + \sum_{i=1}^{r} R_i^{\text{armax}} y(t-i) + e_{\text{resid}}(t) \\ &\quad + \sum_{j=1}^{m} M_j^{\text{armax}} e_{\text{resid}}(t-j) \\ &\quad + \sum_{k=1}^{N_x} \beta_k^{\text{armax}} \mathbf{X}(t,k), \end{aligned} \tag{2}$$

where C^{armax} = a constant coefficient, R_i^{armax} = autoregressive coefficients, M_j^{armax} = moving average coefficients, $e_{\text{resid}}(t)$ = residuals, $y_{\text{armax}}(t)$ = responses, β_k^{armax} = regression coefficients, $\mathbf{X}$ = an explanatory regression matrix in which each column is a time series and $X(t,k)$ denotes an element at the tth row and kth column of the input matrix.

NGARCH(p, q) describes nonlinear time-varying conditional variances and Gaussian residuals $e_{\text{resid}}(t)$. Its mathematical formula is

$$\begin{aligned} \sigma_{\text{ntvcv}}^2(t) &= K^{ng} + \sum_{i=1}^{p} G_i^{ng} \sigma_{\text{ntvcv}}^2(t-i) \\ &\quad + \sum_{j=1}^{q} A_j^{ng} \sigma_{\text{ntvcv}}^2(t-j) \\ &\quad \times \left[\frac{e_{\text{resid}}(t-j)}{\sqrt{\sigma_{\text{ntvcv}}^2(t-j)}} - C_j^{ng} \right]^2 \end{aligned} \tag{3}$$

with constraints

$$\sum_{i=1}^{p} G_i^{ng} + \sum_{j=1}^{q} A_j^{ng} < 1,$$
$$K^{ng} > 0,$$
$$G_i^{ng} \geq 0, \quad i = 1, \ldots, p,$$
$$A_j^{ng} \geq 0, \quad j = 1, \ldots, q,$$

where K^{ng} = a constant coefficient, G_i^{ng} = linear-term coefficients, A_j^{ng} = nonlinear-term coefficients, C_j^{ng} = nonlinear-term thresholds, $\sigma_{\text{ntvcv}}^2(t)$ = a nonlinear time-varying conditional variance and $e_{\text{resid}}(t-j)$ = j-lag Gaussian distributed residual in ARMAX.

In the presence of conditional heteroscedasticity, this composite model can perform ARMAX and NGARCH separately over every period in a time series. For simplicity as employed in [14], it is possible to merge the outputs of ARMAX and NGARCH linearly to attain better results as

$$\begin{aligned} &y_{\text{Composite Model}}(t) \\ &\quad = f(y_{\text{ARMAX}}(t), e_{\text{resid}}(t)) \\ &\quad = Cf_1 \cdot y_{\text{ARMAX}}(t) + Cf_2 \cdot e_{\text{resid}}(t), \end{aligned} \tag{4}$$

where f is defined as a linear function of ARMAX and NGARCH outputs, $y_{\text{ARMAX}}(t)$ and $e_{\text{resid}}(t)$. Cf_1 and Cf_2 in (4) are the coefficients of a linear combination of ARMAX and NGARCH outputs. The resulting residual $e_{\text{resid}}(t)$ at time t is obtained from a product of $\sqrt{\sigma_{\text{ntvcv}}^2(t)}$ in (3) and a normalized random number randn(1) where $0 \leq \text{randn}(1) \leq 1$.

2.3 ANFIS coordinated with NGARCH to improve regression

Two basic aspects of the efficiency and effectiveness of the ARMAX/NGARCH composite model need to be discussed here. First, the conditional mean component of the composite model, ARMAX, is suitable to fit regular or periodic time series, for instance monthly temperature predictions in a local area [15] or seasonal power consumption predictions in a metropolitan area [16]. Strictly speaking, ARMAX cannot fit data sequences very well for irregular or non-periodic time series due to the lack of a dynamic learning mechanism. So, we propose an improved approach, i.e. to replace ARMAX with ANFIS for the conditional mean component of the composite model because ANFIS has its own self-adaptive learning ability to fit irregular or non-periodic time series. Second, when we apply the ANFIS system individually to forecasting a time series and this time series is associated with serial dependence, its performance is considerably impeded by volatility clustering effects. For instance, as shown in Fig. 3, large residual errors have occurred in the ANFIS prediction at sample numbers 9, 12, 21, 24 and 28 among the forecasts of monthly equity volume index futures for the 24 months from Jan. 2001 to Dec. 2002. We therefore incorporate the conditional variance component, NGARCH, into the ANFIS system to help alleviate volatility clustering effects because NGARCH can tackle nonlinear time-varying conditional variances of residuals.

To summarize, based on the relative efficiencies of the available modeling systems and the statistical nature of the reality being modeled, we suggest that ARMAX is substituted by ANFIS in the conditional mean component of the composite model to enhance data fitting of irregular non-periodic time series, with the conditional variance component NGARCH cooperating with the conditional mean to overcome the problem of volatility clustering. This proposed composite model is rewritten as ANFIS/NGARCH. Likewise, a linear combination of ANFIS and NGARCH is recommended so that the ANFIS/NGARCH composite model operates separately and is adapted dynamically to improve the accuracy of irregular non-periodic time series prediction. Formulation of the linear combination [14] is expressed as

$$\begin{aligned} &y_{\text{Proposed Composite Model}}(t) \\ &\quad = g(y_{\text{ANFIS}}(t), e_{\text{resid}}(t)) \\ &\quad = Coef_1 \cdot y_{\text{ANFIS}}(t) + Coef_2 \cdot e_{\text{resid}}(t), \end{aligned} \tag{5}$$

where g is defined as a linear function of the ANFIS and NGARCH outputs, respectively, $e_{\text{resid}}(t)$ and $y_{\text{ANGARCH}}(t)$, while $Coef_1$ and $Coef_2$ are respectively the coefficients of the linear combination of the ANFIS and NGARCH outputs.

Another question is how to determine the coefficients (i.e. $Coef_1$ and $Coef_2$) of the linear combination of ANFIS and NGARCH. The following discussion assists us in choosing the best way to estimate optimal or near-optimal coefficients. Simple regression in statistics is often used to estimate the coefficients of a linear regression. However, it cannot capture the correct behavior of a non-stationary data sequence or a data sequence with very low correlation. This is because a long-term level produced by simple regression cannot express correctly the behavior of a non-stationary data sequence. Further, a data sequence with a very low correlation implies highly nonlinear dynamics such that simple regression is not able to explain adequately a sequence's correlation. An alternative is to consider a nonlinear learning system, e.g. an artificial neural network (ANN), to handle the coefficient estimation. Even though ANNs are frequently used to model nonlinear systems, they may have trouble with over-fitting or under-fitting results due to inappropriate parameters (weights) chosen after training [17]. When the observations are sparse, learning in an ANN may converge to a premature state in which a local optimum is obtained rather than the global optimum. In contrast, quantum-minimized adaptation can outperform an ANN for obtaining globally optimal or near-optimal coefficients when few observations are available for modeling. A novel adaptation mechanism, called quantum minimization (QM) [12], is presented in the next section and will be exploited to search for optimal or near-optimal coefficients $Coef_1$ and $Coef_2$ in (5).

3 Quantum minimization adapting ANFIS/NGARCH

Several methods have been proposed for general constraint optimization [18], yet the lack of an efficient algorithm for quadratic-programming (QP) may turn optimization into a problem with NP complexity. Even though artificial neural networks can be applied to handling optimization, ANNs can not guarantee efficiency when dealing with the NP complexity problem efficiently. In recent years, quantum computing (QC) [19] has offered a promising paradigm for solution of complex problems [20] such as large-number factorization and exhaustive search. A basic feature of QC involves a digital representation of processed information (including both the free parameters and the cost-function computation). Optimization processes based on QC have to scan exhaustively the set of possible bit configurations in the search space.

3.1 Quantum-based optimization

Based on the possibility of superposed states, it is possible that QC can outperform classical computing paradigms [19]. In the quantum-based representation, a quantum bit (a "qbit"), ψ, is allowed to be both state '0' and '1' simultaneously. Each state is characterized by a complex number, γ_i, indicating the probability amplitude of the state:

$$|\psi\rangle = \sum_{i=1}^{n} \gamma_i |\psi_i\rangle, \tag{6}$$

$$\sum_{i=1}^{n} \|\gamma_i\|^2 = 1, \tag{7}$$

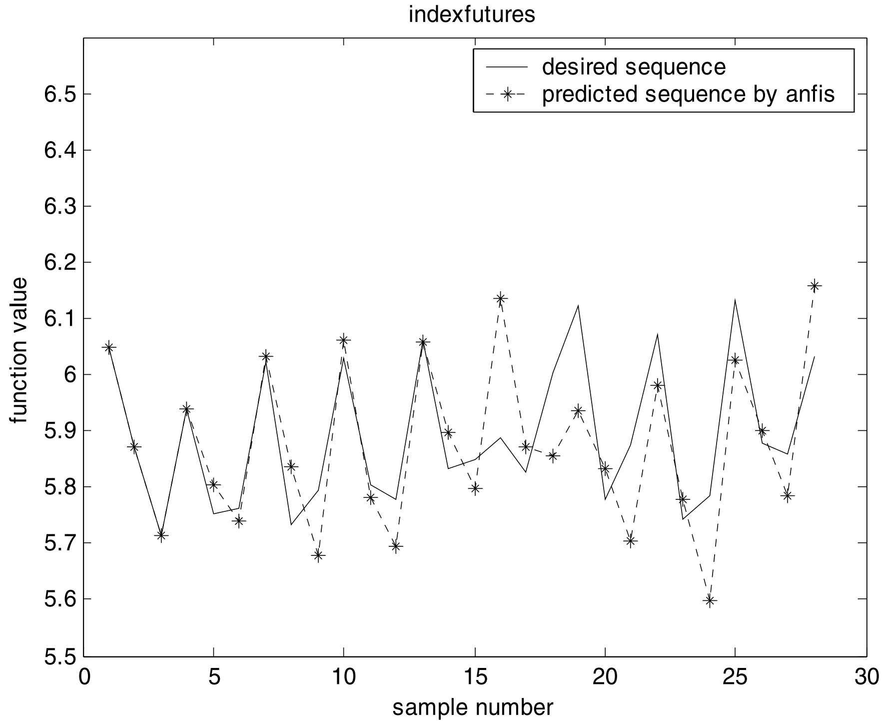


Fig. 3 Forecasts using ANFIS approach exhibit some big residual errors

where "ket" notation $|\bullet\rangle$ has been used for the qbits ψ_i; the probability amplitudes, γ_i, have to satisfy (7).

Thus, for a system with n state bits, a classical computer requires $N = 2^n$ locations for storing all possible configurations, whereas a quantum computer uses just n qbits to represent the whole system state. The basic advantage deriving from superposed states means a quantum computer can explore all of the cost configurations in a single computational run. This is usually achieved by a two-step process: first, an initial state prepared suitably as

$$|\psi_0\rangle = \frac{1}{\sqrt{2^n}} \bigotimes_{i=1}^{n} (|0\rangle + |1\rangle), \tag{8}$$

where $\otimes$ denotes the state direct product [19]; thus $|\psi\rangle$ comprehends all possible states, which are equiprobable. Secondly, feeding the initial state $|\psi\rangle$ to the cost-function algorithm supports an exhaustive scanning of the cost space, thus obtaining a superposition of all possible cost values. By contrast, a classical computer would face an exponential computational overhead. Grover's algorithm [21] is one of the best-known QC techniques proposed so far: it tackles the (NP-complete) problem of searching an input string within an unsorted database. For an input string including n bits and $N = 2^n$ possible states, this database searching problem can obviously be solved in $O(\log N)$ probes if the database is sorted, but no classical algorithm can succeed in the general case with probability better than, say 50%, without probing more than half the entries of the database [12]. By Grover's algorithm [21], however, the number of repetitions grows as $O(\sqrt{N})$. Thus Grover's method does not break the NP-completeness barrier, yet it has represented a popular basis for a large variety of algorithms. In practice, a quantum method for minimization is described in [12]. The number, R, of repetitions for convergence of that algorithm is given by:

$$R = 22.5\sqrt{N} + 1.4\,\mathrm{lg}^2 N \cong 22.5\sqrt{N}, \tag{9}$$

where lg indicates the binary logarithm. Theory shows that a single run of the minimization algorithm [12] obtains a valid solution with probability of at least $1/2$. Therefore, to increase the success probability one merely applies the basic algorithm in a series of $l > 1$ different runs. With this approach, the total number of repetitions, i.e., the computational cost for the quantum machine, is denoted by $R^{(l)}$ and the associated probability of success becomes $P_{qc}^{(l)}$.

$$R^{(l)} = lR, \tag{10}$$

$$P_{qc}^{(l)} \geq 1 - \frac{1}{2^l}. \tag{11}$$

It is seen, then, that the quantum QC techniques discussed above increase their relative effectiveness as the difficulty of the problem increases.

3.2 Quantum minimization (QM)

Quantum minimization (QM) with optimization success probability of at least $1/2$ just in a single probe for an unsorted database can be realized by a quantum minimum searching algorithm [12]. A quantum exponential searching algorithm [22] is called by a quantum minimum searching algorithm as a subroutine to serve in a fast database searching engine.

3.2.1 *Quantum exponential searching algorithm*

Based on the quantum exponential searching algorithm of [22], where the search problem is to find the index i such that $T_{db}[i] = x$, we shall modify their algorithm for finding a solution even though the number t of solutions is known. For simplicity, we assume firstly that $1 \le t \le 3N/4$.

Step 1: Initialize $m = 1$ and set $\lambda = 8/7$. (Any value of λ strictly between 1 and 4/3 will do.)
Step 2: Choose an integer j uniformly at random such that $0 \le j < m$.
Step 3: Apply j iterations of Grover's algorithm [21] starting from initial state

$$|\Psi_0\rangle = \sum_i \frac{1}{\sqrt{N}} |i\rangle.$$

Step 4: Observe the register: let i be the outcome.
Step 5: If $T_{db}[i] = x$, the problem is solved: exit.
Step 6: Otherwise, set m to $\min(\lambda m, \sqrt{N})$ and go back to Step 2.

3.2.2 *Quantum minimum searching algorithm*

We secondly use the minimum searching algorithm of [12] in which the minimum searching problem is to find the index i such that $T_{db}[i]$ is minimum where $T_{db}[0, \dots, N-1]$ is an unsorted table of N items, each holding a value from an ordered set.

Step 1: Choose threshold index $0 \le i \le N-1$ uniformly at random.
Step 2: Repeat the following (2a, 2b, and 2c) and interrupt it when the total running time is more than $22.5\sqrt{N} + 1.4 \lg^2 N$. Then go to stage 3.
  (a) Initialize the memory as $\sum_j \frac{1}{\sqrt{N}} |j\rangle|i\rangle$. Mark every item j for which $T_{db}[j] < T_{db}[i]$.
  (b) Apply the quantum exponential searching algorithm of [22].
  (c) Observe the first register: let i' be the outcome. If $T_{db}[i'] < T_{db}[i]$, then set threshold index i to i'.
Step 3: Return i.

This process is repeated until the probability that the threshold index selects the minimum becomes sufficiently large.

3.3 ANFIS/NGARCH adapted by QM dealing with non-stationary signal

For a non-stationary or highly nonlinear time series such as a random walk [2], it is very difficult to analyze the series dependence characteristics of the data sequence by a traditional autocorrelation function (ACF) or partial autocorrelation function (PACF). This is because the covariance of the data sequence is not constant and actually varies with time.

3.3.1 *Information of signal deviation*

For non-stationary signals, the signal difference (or deviation) of (12) can provide precious information about the short-run dynamics of the currently applied data sequence. A signal deviation $\delta o(k)$ is defined as the backward difference between two consecutively adjacent observations, $o(k)$ and $o(k-1)$, as

$$\delta o(k) = o(k) - o(k-1). \tag{12}$$

The following profiles two aspects associated with signal difference. First, signal deviation is intended to transform two raw signals into a difference value (i.e. signal deviation) that can maintain a high signal-to-noise ratio (SNR) [23] whether or not disturbance exists in the raw signal. This implies that signal deviation makes the system immune to the noise interference which disturbs raw signals [24]. Second, based on transformation, the transformed signal turns out to be a signal sequence with weak stationary property (i.e. approximately constant covariance among the signals) [25]. In other words, we treat these signal deviations as a stationary-like time series on which traditional time series analysis might work. However, traditional analysis is still invalid if the effect of volatility clustering has occurred during the stationary-like time series. Moreover, when modeling a non-periodic short-term regression based on insufficient signal deviations, traditional analysis does not work either. Therefore, the ANFIS/NGARCH composite model adapted by quantum minimization (denoted as the QM-AFNG system) as shown in Fig. 4 can deal with a short-term time series with volatility clustering. This is done by modeling the transformed signal deviations as short-term regression and we subsequently applying trained regression to forecast the future value or estimating the function value.

3.3.2 *QM-AFNG forecasting based on signal deviation*

Single-step-look-ahead prediction, as shown in Fig. 5, can be arranged by adding the most recent predicted signal deviation $\delta\hat{o}(k+1)$ of (13) to the observed current output $o(k)$. The summation results in a predicted output $\hat{o}(k+1)$ at the next period as expressed in (14) [26]. The function h in (13) represents a predictor that includes a data preprocessing

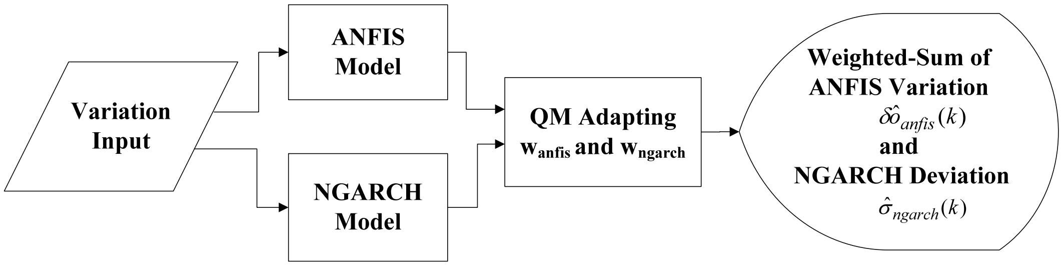


Fig. 4 Diagram of QM adapting ANFIS/NGARCH outputs (denoted as QM-AFNG)

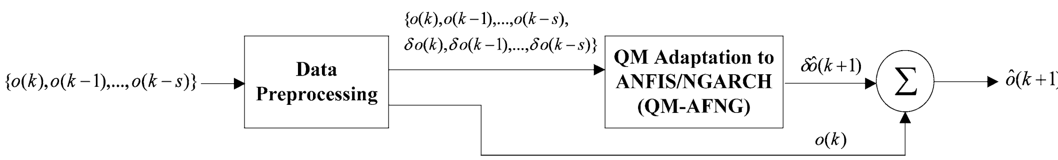


Fig. 5 Prediction using QM-AFNG system

unit, a QM-AFNG system and a summation unit, as shown in Fig. 5. A data preprocessing unit is used to calculate signal deviations of (12) as

$$\delta\hat{o}(k+1) = h(o(k), o(k-1), \dots, o(k-s), \delta o(k), \delta o(k-1), \dots, \delta o(k-s)), \tag{13}$$

$$\hat{o}(k+1) = o(k) + \delta\hat{o}(k+1). \tag{14}$$

Let's turn back and examine again the QM-AFNG system as shown in Fig. 4. In order to construct an ANFIS-based prediction, the most recent predicted deviation $\delta\hat{o}(k+1)$ at next period is assigned as the output of the QM-AFNG system. As shown in Fig. 5, the most recent observations and their deviations, $\{o(k), o(k-1), \dots, o(k-s), \delta o(k), \delta o(k-1), \dots, \delta o(k-s)\}$, have been specified as inputs of the QM-AFNG system. The square-root of nonlinear conditional heteroscedasticity $\hat{\sigma}_{\text{ngarch}}(k)$, not $e_{\text{resid}}(t)$, is derived from the variation sequence of true observations $\{\delta o_{\text{anfis}}(k), \delta o_{\text{anfis}}(k-1), \delta o_{\text{anfis}}(k-2), \delta o_{\text{anfis}}(k-3), \dots\}$. Based on the QM-AFNG structure, one can form the function p of the ANFIS output, $\delta\hat{o}_{\text{anfis}}(k+1)$, and the square-root of NGARCH's output, $\hat{\sigma}_{\text{ngarch}}(k+1)$, as presented below and shown in Fig. 4.

$$\delta\hat{o}_{\text{qm-afng}}(k+1) = p(\delta\hat{o}_{\text{anfis}}(k+1), \hat{\sigma}_{\text{ngarch}}(k+1)) \tag{15}$$

A weighted-average function is assumed to combine both $\delta\hat{o}_{\text{anfis}}(k+1)$ and $\hat{\sigma}_{\text{ngarch}}(k+1)$ to attain a near-optimal result $\delta\hat{o}_{\text{qm-afng}}(k+1)$.

$$\begin{aligned} \delta\hat{o}_{\text{qm-afng}}(k+1) &= w_{\text{anfis}} \cdot \delta\hat{o}_{\text{anfis}}(k+1) \\ &\quad + w_{\text{ngarch}} \cdot \hat{\sigma}_{\text{ngarch}}(k+1) \\ \text{s.t.} \quad & w_{\text{anfis}} + w_{\text{ngarch}} = 1. \end{aligned} \tag{16}$$

Here, the linear combination of two nonlinear functions, $\delta\hat{o}_{\text{anfis}}(k+1)$ and $\hat{\sigma}_{\text{ngarch}}(k+1)$, can also optimally approximate an unknown nonlinear target $\delta\hat{o}_{\text{qm-afng}}(k+1)$. The reason for the approach of (16) is that individual nonlinear function implemented by soft-computing [27] is fast and effective, speeding convergence and reducing computational time.

3.3.3 *Weight-seeking by quantum minimization*

Let $W_{\text{afng}} = [w_{\text{anfis}} \; w_{\text{ngarch}}]^T$ denote a weight-vector of w_{anfis} and w_{ngarch}. A digital cost-function (DCF) [28] is defined as

$$\text{DCF} = \frac{\| W_{\text{afng}} \|^2}{2} + K_{\text{DCF}} \cdot \sum_{k=0}^{l-1} \| y(k+1) - y(k) - o(k) - \delta\hat{o}_{\text{qm-afng}}(k+1) \|^2, \tag{17}$$

K_{DCF}: a regulation coefficient which can be used for measuring accuracy when the cost is minimized. We are hereby seeking a promising searching method, e.g. the quantum searching algorithm, for an optimization search in an unsorted weight-space to look for the appropriate weights in (16) to minimize the DCF of (17) so as to obtain better accuracy for non-periodic short-term time series forecasting.

The reason why quantum approaches are appealing when applied to searching problems is that quantum computing involves a digital representation of processed information such that it is feasible to scan exhaustively the set of possible bit configurations in the search space. Thus, the quantum minimization mentioned above is employed for adapting the weights w_{anfis} and w_{ngarch} for forecasting $\delta\hat{o}_{\text{anfis}}(k+1)$ and $\hat{\sigma}_{\delta o}(k+1)$, respectively, as per (16). In other words, applying quantum computing to the optimization search in the weight-space first requires one to express the weights in a digital representation which is encoded as 8-bit values and which varies in the range $[-100, 100]$ for 256 discrete values. The set of digital weights to be optimized are stored as a set of qbits that are prepared during an initial, equiprobable superposition. Then, the initial state is fed to the cost-function supports for an exhaustive scanning of the weight-space.

4 Experimental results and discussions

4.1 Criteria for measuring accuracy

A critical point in time series prediction is what criterion can be used to measure the accuracy of the predicted results. Because of the wide applications of time series prediction, the question of whether accuracy is "good enough" is largely dependent on user-specified criteria. Measurement of forecasting performance is highly dependent on how rigidly the criteria are specified for measurement of the degree of accuracy. The question of reasonable accuracy for a time series forecast is commonly evaluated by the use of four well-known criteria [29]: (a) mean absolute deviation (MAD); (b) mean absolute percent error (MAPE); (c) mean squared error (MSE); (d) Theil'U inequality coefficient (Theil'U). Specifically, they are defined as

$$\mathrm{MAD} = \frac{\sum_{t=1}^{l} |y_{t_c+t} - \hat{y}_{t_c+t}|}{l}, \tag{18}$$

$$\mathrm{MAPE} = \frac{100}{l} \sum_{t=1}^{l} \left| \frac{y_{t_c+t} - \hat{y}_{t_c+t}}{y_{t_c+t}} \right| \%, \tag{19}$$

$$\mathrm{MSE} = \frac{\sum_{t=1}^{l} (y_{t_c+t} - \hat{y}_{t_c+t})^2}{l}, \tag{20}$$

$$\mathrm{Theil'U} = \sqrt{\frac{\mathrm{MSE}}{\mathrm{MS}}} = \sqrt{\frac{\sum_{t=1}^{l} (y_{t_c+t} - \hat{y}_{t_c+t})^2 / l}{\sum_{t=1}^{l} y_{t_c+t}^2 / l}}, \tag{21}$$

where l = the number of periods in forecasting, t_c = the current period, y_{t_c+t} = a desired value at the $t_c + t$th period and $\hat{y}_{t_c+t}$ = a predicted value at the $t_c + t$th period.

4.2 Two experiments and their verifications

As shown in Figs. 6 to 11, the forecasting abilities of our proposed method and several alternative methods are compared in experiments wherein each predictive methodology is applied to two actual historical sets of time series data: (i) the forecast of international stock price indices; (ii) the forecast of equity index futures and options. The alternative methods used are grey model (GM), auto-regressive moving-average (ARMA), back-propagation neural network (BPNN), ARMA/NGARCH composite model (ARMAXNG), adaptive neuro-fuzzy inference system (ANFIS), and the ANFIS/NGARCH composite model adapted by quantum minimization (QM-AFNG). Single-step-look-ahead prediction methodology is employed in all experiments. In single-step-look-ahead design, a small number of the most recent observed data are collected as a sliding window (i.e. data queue) for modeling an intermediate predictor to predict the next period output. Once the next period's sampled datum is obtained, we drop a datum at the bottom of the data queue and add the most recent sampled datum into the data queue at the top position, thereby forming the new data queue used for the next prediction. This process continues until the task is terminated. To simplify comparison of the tested methods as plotted curves, only the three most representatives of our tested models are shown in the figures. Thus GM, ARMA and the proposed QM-AFNG are illustrated in Figs. 6 to 11, where "$\bullet$" represents the sequential output of the GM prediction, "$\circ$" represents the sequential output of the ARMA prediction and "$- * -$" represents the sequential output of the QM-AFNG prediction. All six methods, however, are compared for goodness-of-fit in Tables 1 to 8.

First, as shown in Figs. 6 to 9, the forecast of international stock price indices of four markets (New York Dow-Jones Industrials Index, London FTSE-100 Index, Tokyo Nikkei Index and Taipei Taiex Index) [30] have been tested with 36 points, 36 points and 12 points taken from the most recent historical data as training, testing and validating samples, respectively. In addition, comparative performance is obtained by comparing the actual sampled values and the predicted results of international stock price monthly indices over 48 months from Jan. 2002 to Dec. 2005 for (a) mean absolute deviation (MAD), (b) mean absolute percent error (MAPE) $\times 100$, (c) mean squared error (MSE) (unit $= 10^5$) and (d) Theil'U inequality coefficient (Theil'U). Forecasting performance of all six methods is summarized in Tables 1 to 4, showing QM-AFNG obtains the best prediction results. The goodness of fit of QM-AFNG prediction modeling for the four markets is tested by the Ljung-Box Q-test [31] with p-values of 0.5082, 0.3239, 0.4751 and 0.3702, where each p-value is greater than the level of significance (0.05).

Figures 10 and 11 show comparative forecasts of two typhoon moving paths (Nari typhoon for September 6–19, 2001 and Mindulle typhoon for June 28–July 3, 2004) [32]. Likewise, modeling for the second experiment takes training, testing and validating samples from the most recent historical data by 14 points, 14 points and 7 points, respectively. Performance evaluation is again made by comparison of the actual and predicted values for MAD, MAPE, MSE and Theil'U. Tables 5 to 8 summarize prediction performance of our alternative methods and show that QM-AFNG achieves

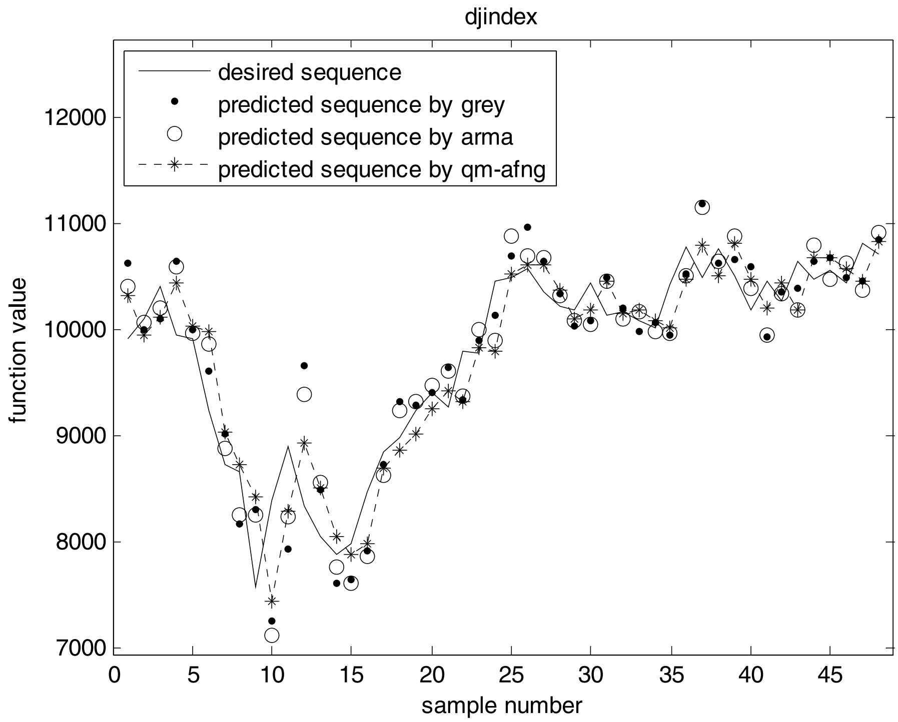


Fig. 6 Forecasts of monthly New York D.J. industry index

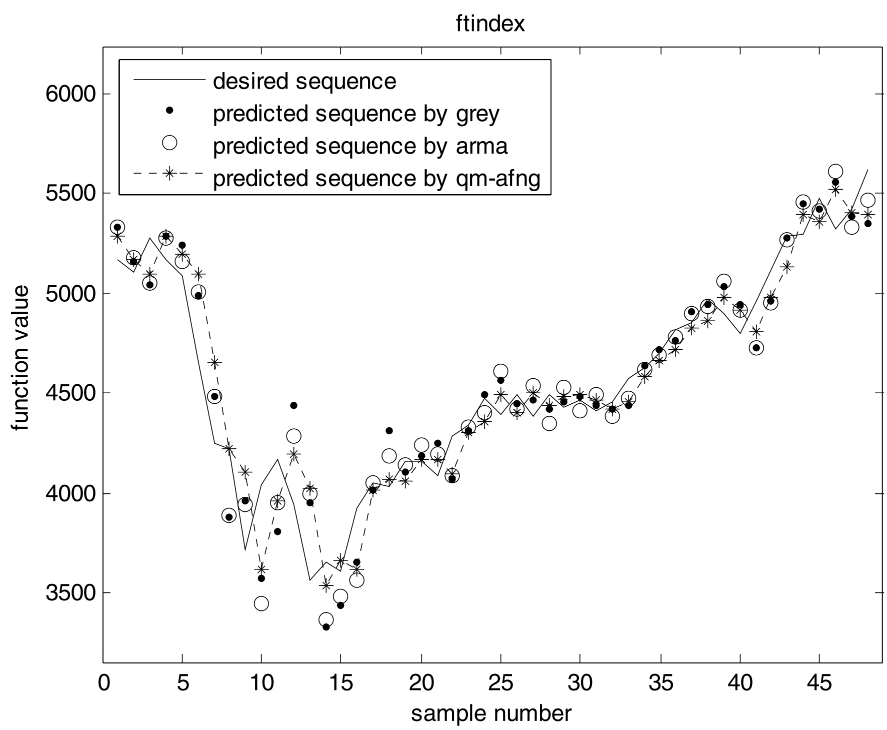


Fig. 7 Forecasts of monthly London FTSE-100 index

superior results. The goodness of fit of QM-AFNG prediction modeling for the futures and options data is also tested by Ljung-Box Q-test with p-values of 0.1948 and 0.2375, in which each p-value is greater than the level of significance (0.05).

4.3 Discussion

Computational complexity and predictive accuracy are two critical issues which must be considered when modeling complex time-series. Regarding the computational cost of

Fig. 8 Forecasts of monthly Tokyo Nikkei index

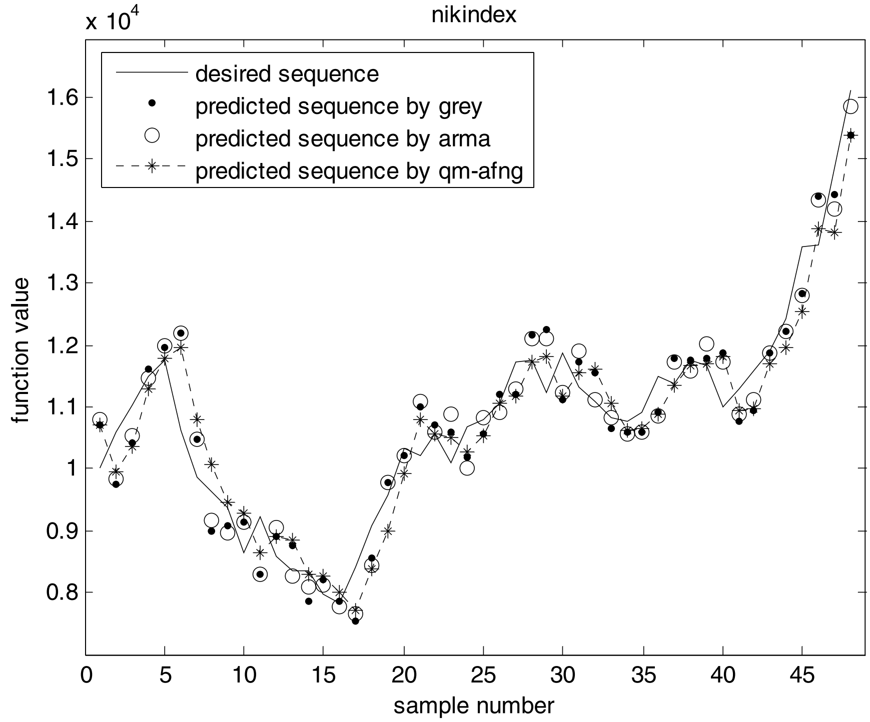
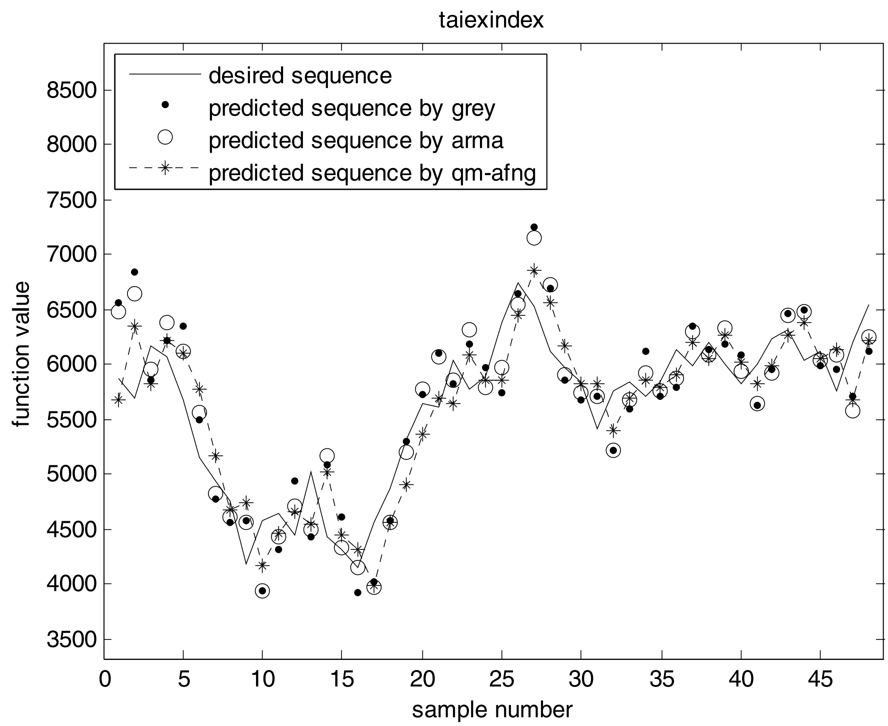


Fig. 9 Forecasts of monthly Taipei Taiex index



the various models tested, the proposed method requires the greatest computational effort to find the optimal ANFIS and NGARCH combination. In fact, the amount of time required depends on the computational complexity of the quantum minimum searching algorithm. Consequently, quantum minimum searching costs around $O(\sqrt{N}+ld^2(N))$ for searching an unsorted database where an input string includes n bits and $N=2^n$ represents all possible states.

As indicated, (5) represents a plane in a three-dimensional space where the two coefficients $Coef_1$ and $Coef_2$ can be viewed as slopes associated with the plane. The second coefficient $Coef_2$ normally attains a very small real value, e.g.

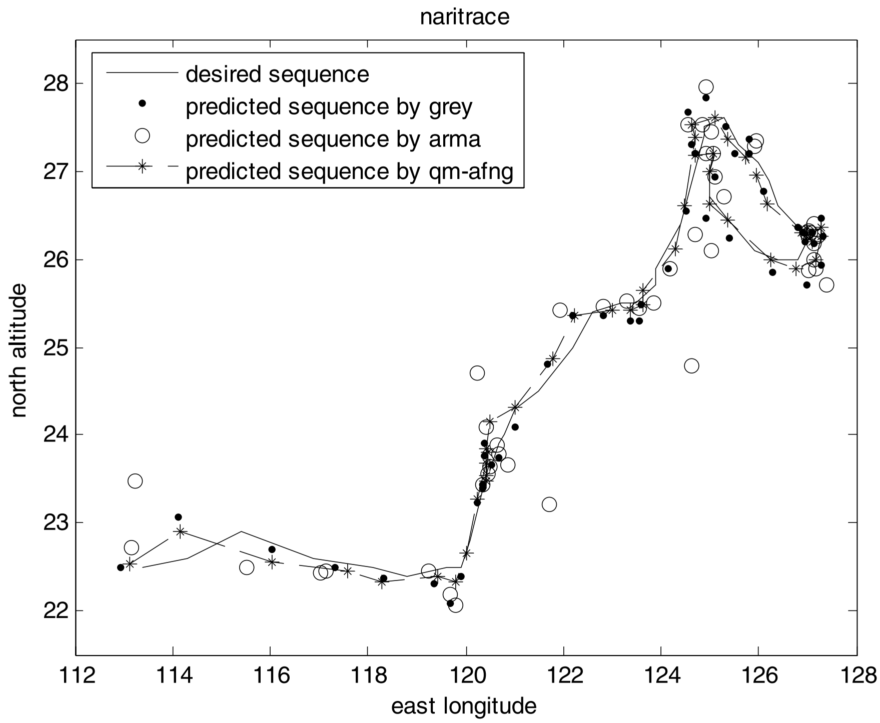


Fig. 10 Forecasts of Nari typhoon moving path

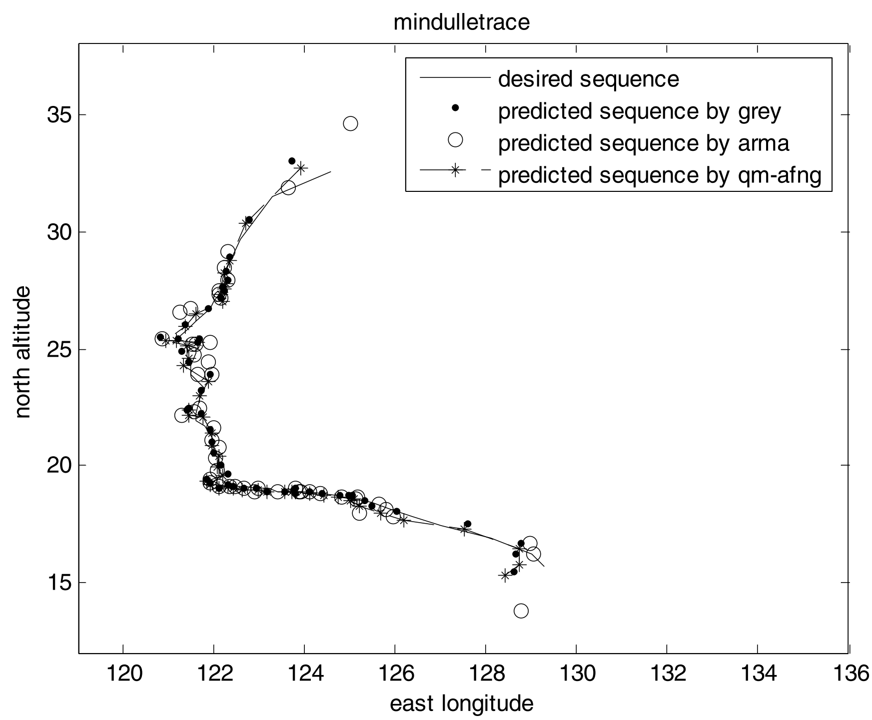


Fig. 11 Forecasts of Mindulle typhoon moving path

0.0528, which is used to adjust the conditional standard deviation of the residuals. In consequence, the first coefficient $Coef_1$ gains a real value, e.g. 0.9472, which is always bigger than the second one. Since the predicted outputs are very sensitive to the changes in slopes (i.e. coefficients), the coefficients should be captured as precisely as possible so as not to lose the generality of a set of complex time-series data. This will result in a best-fit model for a complex time-series, improving prediction based on analysis of the series' dynamic properties.

The experimental results have shown that the proposed method produces the most satisfying solutions and thus is an

Table 1 The comparison between different prediction models based on Mean Absolute Deviation (MAD) on international stock price monthly indices

Methods	Mean Absolute Deviation (MAD)					
	New York D.J. Industrials Index	London FTSE-100 Index	Tokyo Nikkei Index	Taipei TAIEX Index	Average	Standard Deviation
GM	340.5970	153.8277	477.2157	355.1361	331.6941	133.4688
ARMA	339.7215	153.7628	439.8190	321.1152	313.6046	118.6282
BPNN	279.1350	134.5064	453.7069	277.5879	286.2341	130.6313
ARMAXNG	320.7695	152.3504	437.0319	317.9291	307.0202	117.0952
ANFIS	284.5725	145.3118	441.5919	296.1719	291.9120	121.0616
QM-AFNG	218.7492	111.3910	406.3961	239.9533	244.1224	121.9447

Table 2 The comparison between different prediction models based on Mean Absolute Percent Error (MAPE) on international stock price monthly indices

Methods	Mean Absolute Percent Error (unit $= 10^{-2}$) (MAPE)					
	New York D.J. Industrials Index	London FTSE-100 Index	Tokyo Nikkei Index	Taipei TAIEX Index	Average	Standard Deviation
GM	3.65	3.54	4.49	6.49	4.54	1.365903
ARMA	3.61	3.53	4.14	5.81	4.27	1.060138
BPNN	2.98	3.06	4.19	5.05	3.82	0.98877
ARMAXNG	3.52	3.50	4.12	5.77	4.23	1.067813
ANFIS	3.06	3.31	4.13	5.40	3.98	1.054214
QM-AFNG	2.51	2.63	3.68	4.27	3.27	0.847600

Table 3 The comparison between different prediction models based on Mean Squared Error (MSE) on international stock price monthly indices

Methods	Mean Squared Error (unit $= 10^5$) (MSE)					
	New York D.J. Industrials Index	London FTSE-100 Index	Tokyo Nikkei Index	Taipei TAIEX Index	Average	Standard Deviation
GM	1.9582	0.40063	3.2209	1.7472	1.8317	1.154842
ARMA	1.8230	0.38832	2.9384	1.4737	1.6559	1.050823
BPNN	1.2652	0.30656	3.0189	1.0461	1.4092	1.148844
ARMAXNG	1.8170	0.38527	2.9193	1.4772	1.6497	1.043777
ANFIS	1.3550	0.38494	2.8912	1.1683	1.4499	1.048775
QM-AFNG	0.9488	0.25307	2.2140	0.7886	1.0511	0.830373

attractive approach to modeling complex time-series. This is because quantum-based minimization applied to global minimum searching has greatly enhanced the optimal tuning of the two coefficients for combining the above-mentioned models, resulted in overcoming both the overshoot and volatility clustering effects at the same time. Results show highly improved predictive accuracy which significantly outperforms such well-known systems as auto-regressive moving-average, back-propagation neural network and single fuzzy inference. These improved results are attributable to improved accuracy of the coefficients when they are obtained by quantum-based searching, due to the superior ability of quantum-based searching to find the best coefficients from a large group of possible states.

Apparently quantum-based searching is definitely capable of handling the task of finding multiple parameters for optimizing highly nonlinear model [33] very well. In order to tackle overshoot predicted results and resolve volatil-

Table 4 The comparison between different prediction models based on Theil'U Inequality Coefficient (Theil'U) on international stock price monthly indices

Methods	Theil'U Inequality Coefficient (Theil'U)					
	New York D.J. Industrials Index	London FTSE-100 Index	Tokyo Nikkei Index	Taipei TAIEX Index	Average	Standard Deviation
GM	0.0435	0.0414	0.0501	0.0721	0.0518	0.014048
ARMA	0.0420	0.0408	0.0479	0.0662	0.0492	0.011734
BPNN	0.0349	0.0362	0.0485	0.0558	0.0439	0.010051
ARMAXNG	0.0409	0.0406	0.0477	0.0663	0.0489	0.012070
ANFIS	0.0362	0.0411	0.0475	0.0590	0.0460	0.009854
QM-AFNG	0.02877	0.0319	0.0389	0.0512	0.0377	0.009951

Table 5 The comparison between different prediction models based on Mean Absolute Deviation (MAD) on futures and options volumes monthly indices of equity products

Methods	Mean Absolute Deviations (MAD)			
	Nari Typhoon Moving Path	Mindulle Typhoon Moving Path	Average	Standard Deviation
GM	0.2053	0.1796	0.1925	0.018173
ARMA	0.3060	0.2197	0.2629	0.061023
BPNN	0.2784	0.6654	0.4719	0.27365
ARMAXNG	0.2772	0.6389	0.4581	0.255761
ANFIS	0.2210	0.5402	0.3806	0.225708
QM-AFNG	0.1884	0.0937	0.1411	0.066963

Table 6 The comparison between different prediction models based on Mean Absolute Percent Error (MAPE) on futures and options volumes monthly indices of equity products

Methods	Mean Absolute Percent Error (MAPE)			
	Nari Typhoon Moving Path	Mindulle Typhoon Moving Path	Average	Standard Deviation
GM	0.0047	0.0043	0.0045	0.000283
ARMA	0.0075	0.0056	0.0066	0.001344
BPNN	0.0059	0.0087	0.0073	0.00198
ARMAXNG	0.0055	0.0085	0.0070	0.002121
ANFIS	0.0050	0.0072	0.0061	0.001556
QM-AFNG	0.0043	0.0028	0.0036	0.001061

ity clustering effect simultaneously, the forecasting system herein is well organized as a least structure to form a linear combination of two distinct functions (models) for simplicity purpose [14] and their combinative coefficients have found optimally by quantum-based searching. Even though only two optimal parameters have been discovered to complete the model via quantum technique, this work has stressed a novel approach in view of introducing the least structure (i.e. the minimum cost) with excellent performance (i.e. the best accuracy) according to the above discussion over computational complexity and predictive accuracy.

5 Conclusions

When using the ANFIS approach alone, volatility clustering degrades the effectiveness and efficiency of time series prediction by inducing large residual error. Dealing with this flaw has become an urgent issue. This study has proposed a method that incorporates a nonlinear generalized autoregressive conditional heteroscedasticity (NGARCH) into an ANFIS approach so as to correct the crucial problem of time-varying conditional variance in residual errors. In other words, we have constructed a linear combination of ANFIS

Table 7 The comparison between different prediction models based on Mean Squared Error (MSE) on futures and options volumes monthly indices of equity products

Methods	Mean Squared Error (MSE)			
	Nari Typhoon Moving Path	Mindulle Typhoon Moving Path	Average	Standard Deviation
GM	0.0648	0.0703	0.0676	0.003889
ARMA	0.2165	0.1614	0.1890	0.038962
BPNN	0.1421	0.1405	0.1413	0.001131
ARMAXNG	0.1497	0.1812	0.1655	0.022274
ANFIS	0.0816	0.1372	0.1094	0.039315
QM-AFNG	0.0599	0.0575	0.0587	0.001697

Table 8 The comparison between different prediction models based on Theil'U Inequality Coefficient (Theil'U) on futures and options volumes monthly indices of equity products

Methods	Theil'U Inequality Coefficient (Theil'U)			
	Nari Typhoon Moving Path	Mindulle Typhoon Moving Path	Average	Standard Deviation
GM	0.0022	0.0023	0.0023	7.07E−05
ARMA	0.0034	0.0028	0.0031	0.000424
BPNN	0.0036	0.0025	0.0031	0.000778
ARMAXNG	0.0038	0.0029	0.0034	0.000636
ANFIS	0.0025	0.0023	0.0024	0.000141
QM-AFNG	0.0018	0.0017	0.0018	7.07E−05

and NGARCH as a time series predictor, in which the coefficients of the combined methods are adapted automatically by quantum minimization; by this adapting method, both the enhancement of generalization and the stability of the proposed predictor are achieved simultaneously. In this manner, large residual error is significantly reduced because the effect of volatility clustering is regulated to a trivial level. Instead of an artificial neural network (ANN), quantum minimization (QM) is employed to adjust and optimize the coefficients for the linear combination of ANFIS and NGARCH, since QM can perform this function with fewer sampled data. Experimental comparison of a range of systems shows that the ANFIS/NGARCH composite model adapted by QM provides superior prediction accuracy and good computation speed for irregular non-periodic short-term time series forecast.

Acknowledgements This work is fully supported by the National Science Council, Taiwan, Republic of China, under grant number NSC 94-2218-E-143-001.

References

1. Chang BR (2003) Advanced hybrid prediction algorithm for non-periodic short-term forecasting. Int J Fuzzy Syst 5(3):151–160
2. Box GEP, Jenkins GM, Reinsel GC (1994) Time series analysis: forecasting & control. Prentice-Hall, New Jersey
3. Chang BR (2004) Tunable free parameters C and epsilon-tube in support vector regression grey prediction model—SVRGM(1,1|C,ε) approach. In: Proc IEEE SMC, pp 2431–2437
4. Castillo O, Melin P (2001) Simulation and forecasting complex economic time series using neural network and fuzzy logic. In: Proc IJCNN, pp 1805–1810
5. Thomson R, Hodgman TC, Yang ZR, Doyle AK (2003) Characterizing proteolytic cleavage site activity using bio-basis function neural networks. Bioinformatics 19(14):1741–1747
6. Deng JL (1982) Control problems of grey system. Syst Control Lett 1(5):288–294
7. Jang J-SR (1993) ANFIS: adaptive-network-based fuzzy inference systems. IEEE Trans Syst Man Cybern 23(3):665–685
8. Wang J-S (2003) An efficient recurrent neuro-fuzzy system for identification and control of dynamic systems. In: Proc IEEE SMC, pp. 2833–2838
9. Neter J, Wasserman W, Kutner MH (1985) Applied linear statistical models, 2nd edn. Homewood, Irwin
10. Bellerslve T (1986) Generalized autoregressive conditional heteroscedasticity. J Econom 31:307–327
11. Gourieroux C (1997) ARCH models and financial applications. Springer, New York
12. Durr C, Hoyer P (1996) A quantum algorithm for finding the minimum, http://arxiv.org/abs/quant-ph/9607014
13. Hamilton JD (1994) Time series analysis. Princeton University Press, New Jersey
14. Pshenichnyj BN, Wilson SS (1994) The linearization method for constrained optimization. Springer, New York
15. Huang YP, Yu TM (1997) The hybrid grey-based models for temperature prediction. IEEE Trans SMC Part B: Cybern 27(2):284–297

16. Mori H, Kosemura N, Ishiguro K, Kondo T (2001) Short-term load forecasting with fuzzy regression tree in power systems. In: Proc IEEE SMC, pp 1948–1953
17. Bose NK, Liang P (1996) Neural network fundamentals with graphs, algorithm, and applications. McGraw-Hill, New York
18. Fletcher R (1987) Practical methods of optimization, 2nd edn. Wiley, New York
19. Nielsen MA, Chuang IL (2000) Quantum computation and quantum information. Cambridge University Press, London
20. Deutsch D, Jozsa R (1992) Rapid solution of problems by quantum computation. Proc R Soc Lond A 439:553–558
21. Grover LK (1996) A fast quantum mechanical algorithm for database search. In: Proc 28th ann ACM symp theory of comp. ACM Press, New York, pp 212–219
22. Boyer M, Brassard G, Hoyer P, Tapp A (1998) Tight bounds on quantum searching. Fortschr Phys 46:493–505
23. Proakis JG (2001) Digital communications, 4th edn. McGraw-Hill, New York
24. Couch LW II (2001) Digital and analog communication systems, 6th edn. Prentice-Hall, Upper Saddle River
25. Priestly MB (1988) Non-linear and non-stationary time series analysis. Academic, New York
26. Chang BR (2005) Hybrid BPNN-weighted grey-CLMS forecasting. J Inf Sci Eng 21(1):209–221
27. Jang J-SR, Sun C-T (1997) Neuro-fuzzy and soft computing: a computational approach to learning and machine intelligence. Prentice-Hall, Upper Sadle River
28. Anguita D, Ridella S, Rivieccio F, Zunino R (2003) Training support vector machines—a quantum-computing perspective. In: Proc IJCNN, pp 1587–1592
29. Diebold FX (1998) Elements of forecasting. South-Western, Cincinnati
30. International Stock Price Index (2002) FIBV focus monthly statistics
31. Ljung GM, Box GEP (1978) On a measure of lack of fit in time series models. Biometrika 65:67–72
32. Central Weather Bureau (2005) Typhoon moving trace inquire. Taipei, Taiwan
33. Imre S, Balazs F (2005) Quantum computing and communications an engineering approach. Wiley, West Sussex

Bao Rong Chang is currently an Associate Professor in the Department of Computer Science and Information Engineering at National Taitung University in Taitung, Taiwan. He completed his BS degree from the Department of Electronic Engineering, Tam Kang University, Taiwan. In 1990, he earned his ME degree from the Department of Electrical Engineering, University of Missouri-Columbia, USA, and his Ph.D. in 1994 at the same University. His current research interests include Intelligent Computation, Embedded System, Computer Network, and Financial Engineering.

Hsiu-Fen Tsai is currently a Senior Lecturer in the Department of International Business at Shu Te University in Kaohsiung, Taiwan. She completed her BA degree from the Department of International Business, National Taiwan University, Taiwan. In 1995, she earned her MBA degree from the Department of Business Administration, National Taiwan University, Taiwan. At present, she is a Ph. D. Candidate in Department of International Business since 2004 at the same University. Her current research interests include Intelligent Analysis of Business Models and Applications of Strategy Management.